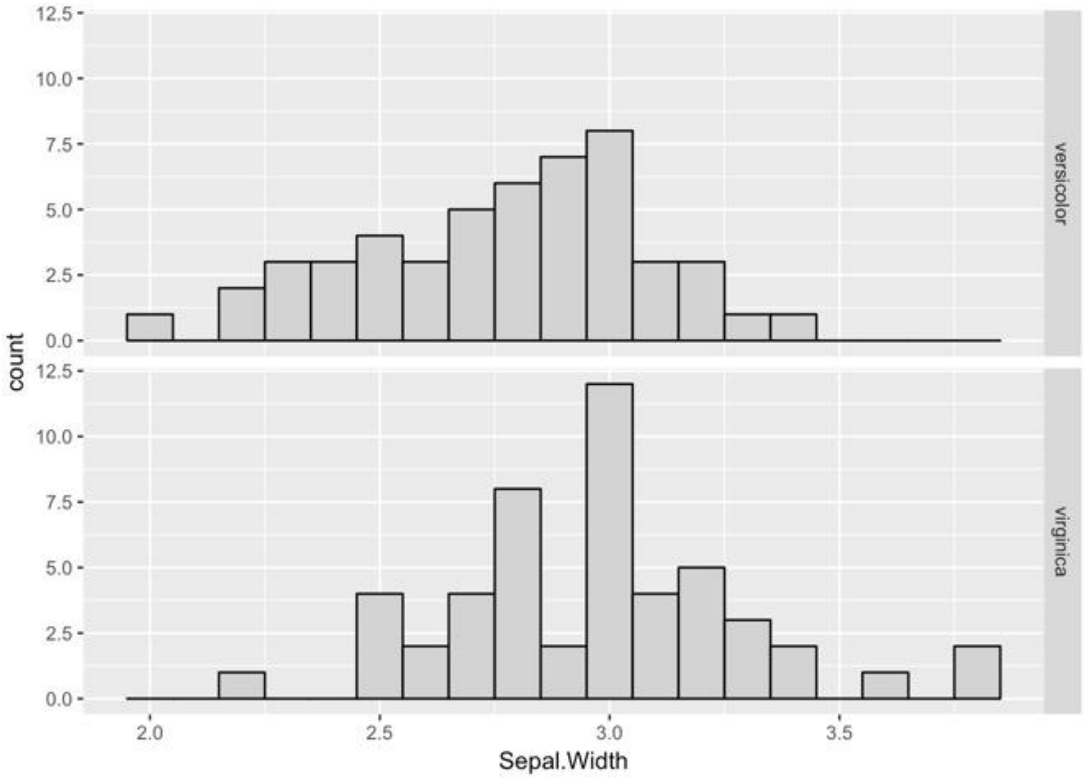
Stats M2

WANG Ying （21406554）

Exercise 2

http://manganese.lab.parisdescartes.fr:18888/files/assignment_1-exercise_2.html

Exercise 3



- **Hypothesis A**: The virginica and versicolor iris species are the same in terms of sepal width.
- **Hypothesis B**: The virginica and versicolor iris species are different in terms of sepal width.

Question 3c:

Explication of these Hypothesis:

These two hypothesis only contain one dependent variable, but not independent variable with which the dependent variable's behavior is explained. The only variable is the sepal width.And the variable is an ordinal variable which can be ranked from low to high.

Hypothesis A says that the sepal width does not play a role in distinguishing the virginica and versicolor iris species.

Whereas Hypothesis B says that the sepal width is the variable between the

virginica and versicolor.

Question 3d:

I am in favor of the Hypothesis B for these reasons:

1 Visually the two histograms show a very different distribution in sepal width:

(1)Among the versicolor specie, there is no a huge difference; but for virginica, the variation of sepal width can be very huge.

(2) The histogram for versicolor is prettty continuous, but there are some vide zones in the histogram for virginica.

2 If we look into details, we can find some datas are not in either of these two spicies, for example:

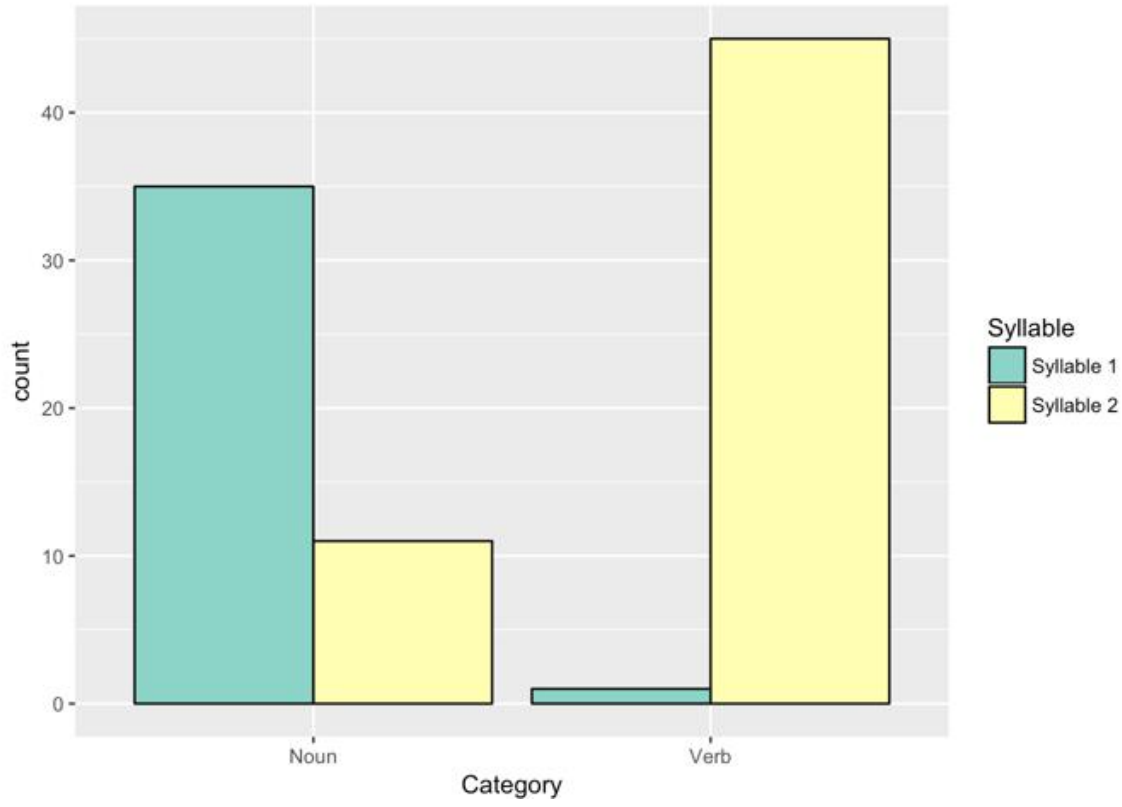
(1) The minimum and maximum width are not same for these two spicies:

Versicolor: minimum < 2.0 ; maximum 3.25 ~ 3.5

Virginica: minimum 2.125 ~ 2.25 ; maximum > 3.75

(2) Among the 23 samples of virginica, whose sepal width is 3.0 is 11, almost half of the total number; but for virginica, only 8 of them.

Exercise 4



- **Hypothesis A**: **Permit (noun)** and **permit (verb)** are the same in terms of their stress.
- **Hypothesis B**: **Permit (noun)** and **permit (verb)** are different in terms of their stress.

Question 4a:

Explication of these Hypothesis:

These two hypothesis only contain one dependent variable, but not independent variable with which the dependent variable's behavior is explained. The only variable is the placement of the stress in the word.But here the variable is not ordinal, but binary since there are only two syllables in the word.

Hypothesis A says that the permit (noun) and the permit(verb) share the same placement of the stress.

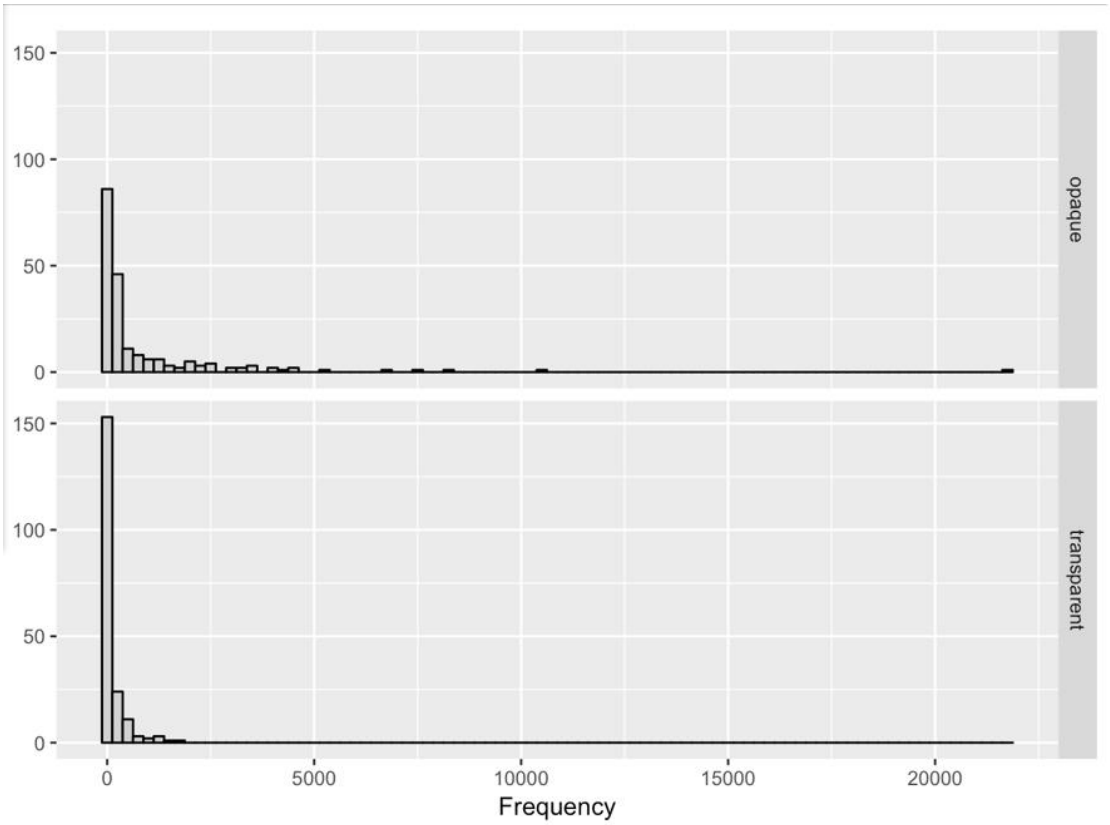
Whereas Hypothesis B says that the permit (noun) and the permit(verb) can be distinguished from the different placement of the stress.

Question 4b:

I am in favor of the Hypothesis B for these reasons:

Among the total number of 46, the placement of the stress on the first syllable for permit (noun) gets 35 and the second syllable only 11. Also, the difference for the permit(verb) is extremely considerable since the placement on the first syllable takes almost zero.

Exercise 5



Question 5a:

The independent variable is the ratio variable FREQUENCY , which can take on all sorts of positive integer values.

The dependent variable is the nominal variable or binary variable, which can be either OPAQUE or TRANSPARENT.

Question 5b:

- **Hypothesis A**: Semantically transparent and opaque *ver-* verbs are the same in terms of their frequency.
- **Hypothesis B**: Semantically transparent and opaque *ver-* verbs are different in terms of their frequency.

I am in favor of the Hypothesis B for the reason that the frequency differ significantly from one and another: in the first histogram, we can even see the frequency between 2500 and 25000, which is not found for the transparent ver-verbs.